Journal of Cyber Policy

ISSN: 2373-8871 (Print) 2373-8898 (Online) Journal homepage: http://www.tandfonline.com/loi/rcyb20

Towards estimating the untapped potential: a global malicious DDoS mean capacity estimate

Eireann Leverett & Aaron Kaplan

To cite this article: Eireann Leverett & Aaron Kaplan (2017) Towards estimating the untapped potential: a global malicious DDoS mean capacity estimate, Journal of Cyber Policy, 2:2, 195-208, DOI: 10.1080/23738871.2017.1362020

To link to this article: http://dx.doi.org/10.1080/23738871.2017.1362020

JOURNAL OF CYBER POLICY, 2017
VOL. 2, NO. 2, 195–208
https://doi.org/10.1080/23738871.2017.1362020

Routledge
Taylor & Francis Group

Towards estimating the untapped potential: a global malicious DDoS mean capacity estimate

Eireann Leverett [a] and Aaron Kaplan[b]

[a]Privacy International, London, UK; [b]Concinnity Risks and CERT, Vienna, Austria

ABSTRACT

What is the malicious reflected distributed denial of service (rDDoS) mean potential of the internet? The authors have been using data from the openNTP project[1] which measures the number of reflectors on the internet since 2014 until now, and completed a graph that roughly estimates a lower boundary for global rDDoS mean potential across four internet protocols (IPs); SSDP, NTP, SNMP and open recursive DNS. By summing these values, and adjusting for average uplink capacity from reflectors, we come to a single number: 108.49 Tb/s as an estimate of rDDoS magnitude potential across IPv4. Tracking this number over time can give us insights into global remediation and clean-up efforts and where to invest our resources in when battling of rDDoS attacks. This paper demonstrates that the upstream throughput is the main contributing, measurable limiting factor for a volumetric rDDoS attack. The largest DDoS event reported by a single target (Dyn in 2016) was 1.2 Tb/s.[2] In contrast, our lower estimate for the global attack potential is *two* orders of magnitude larger than the Dyn attack.

The key contribution is an extensible methodology for measuring global potential for rDDoS attacks, with surprising policy implications.

ARTICLE HISTORY
Received 18 April 2017
Revised 13 July 2017
Accepted 27 July 2017

KEYWORDS
IoT; DDoS; reflective DDoS; volumetric DDoS; amplification attacks; reflectors

Definitions

DDoS	distributed denial of service attack. Specifically, this paper only addresses so-called UDP-based (reflective) volumetric DDoS attacks. See the section on Introduction to DDoS
ASN	autonomous system number. Roughly equivalent to a unique number per internet service provider (ISP)
NTP	network time protocol.[3] A protocol for precise timekeeping on the internet. An essential service on the internet
DNS	domain name service protocol.[4] A protocol on the internet which maps domain names to IP addresses. Essential for the workings of the internet
SSDP	simple service discovery protocol.[5] A non-essential service of the internet, usually deployed on home routers
SNMP	simple network management protocol.[6] An essential service on the internet. It should be deployed on internal IP address space; however, often SNMP is deployed on public IP address space and hence publicly visible
UDP	user datagram protocol
TCP	transmission control protocol
IoT	internet of things

CONTACT Eireann Leverett eireann.leverett@cantab.net

Tb/s	Terabits/s (10^{12} bits/s)
Amplifier	We will use this term synonymously with reflector
AF	amplification factor

Introduction to DDoS and IoT

Due to the continued growth of e-commerce and the internet, distributed denial of service (DDoS) attacks have become a hot policy issue at the international level. The nascency of the unprotected internet of things (IoT) devices also has policy people concerned about the effect of those devices on internet health and on traffic patterns.

Metrics are an important part of triage and remediation of harms, both on and off the internet: 'What we can't measure, we can't improve'.

In this paper we focus on the specific harm of DDoS, and introduce new measurements and metrics for consideration to reduce that harm-potential. This paper should consequently be of interest to DDoS researchers, economists, policy-makers and DDoS cyber-insurance providers alike.

The IoT is expected to flood the internet with cheap, insecure, non-updateable and unpatchable devices, many of which may and do contribute to DDoS attacks. While one could capture this growth only by measuring events over time, the authors have instead taken the approach of measuring the contributors who make those events possible.

Thus we conclude with a single number (108.49 Tb/s) that we believe is a usable estimate of DDoS potential on the internet – based on the measurement of four protocols which are commonly used in DDoS attacks. We realise that there may be other protocols and larger amplifications yet to come. However, that too will be limited by available throughput, as we discuss below. The main contribution of the paper is not the number itself (which will no doubt change over time), but rather the methodology proving that it *is* possible to calculate the global potential. Our research makes the novel finding that it is upstream throughput which is the key constraining factor and revealing enabler of reflective DDoS attacks, and not the number of reflectors as previously assumed. We have reframed the problem as a max-flow problem, using the max-flow of the edge of the network to establish a baseline.

A key focus for remediation should therefore be on larger throughput organisations and data centres in developed countries which also coincidentally often have the resources and skills necessary to implement existing good practices.

Types of DDoS

Not all DDoS attacks are created the same way, or have the same types of impacts at the system level. Indeed they have been exhaustively decomposed into different types and taxonomies, which we will not reproduce here (Asosheh and Ramezani 2008; Barabási 2009; Mirkovic and Reiher 2004).

For the purpose of this paper we are focused exclusively on volumetric and more specific user datagram protocol (UDP) reflected attacks, though we will describe the other types for those new to DDoS policy below.

Transmission control protocol (TCP) connection

Commonly known as SYN flooding, this method of attacking a service involves beginning the opening sequence of a TCP handshake but timing out on further interactions with the TCP session. This means the server managing TCP connections receives many 'Hello' (SYN) messages, but gets no response to 'What can I do for you?', thus leading to a great number of open sessions without any further progression of the TCP session. Over time this exhausts the service of resources and degrades performance.

Fragmentation

Messages passing between computers have size limits, and if you need to send a message larger than this limit, then you need to break it into more pieces and reassemble them at the other end. To do so, you also send some instructions, and a kind of 'some assembly required' message. Let's explore for a moment what would happen if you lie about the messages sizes and how they are reassembled.

As an analogy this is akin to sending a child a box with random model pieces taken from many sets, and a set of instructions. How long would they spend trying to build a model that was impossible to build? With this kind of attack, that is precisely what the attacker does to the computer receiving the messages, who unlike the child in the analogy will always spend the same amount of time realising the messages cannot be reassembled, no matter how often the task is re-attempted.

Computers expect the messages to be sensible and for the instructions to eventually lead to a correctly formed meaningful piece of content. When attackers make unsolvable puzzles, the machines exhaust their memory and processing power in an attempt to reassemble the packets.

Let us move on to other types of DDoS, before we show how we achieved our surprising and counter-intuitive result.

Application

Each web application has a limit on the resources available to it, or the speed at which they can be delivered. These might be for example memory, processing power, disk space and bandwidth. For simplicity, let us refer to these loosely as 'capacity', and the speed at which more capacity can be delivered as 'surge capacity'.

A networked application can be fed inputs that exceed its capacity or surge capacity. This degrades or eliminates the service, fulfilling the attacker's desire to perform a DDoS.

As an illustrative example, imagine a website for credit card payments. While transactions using the card might be highly provisioned with respect to capacity, the system might require a lot more capacity to process new applications. If an attacker overwhelmed the system with new applications, then even the processing of card transactions might degrade or halt.

Volumetric

Lastly, we come to reflected volumetric attacks, which is our primary interest in this paper. In this kind of attack the sheer bandwidth of the messages overwhelms the server or its network bandwidth.

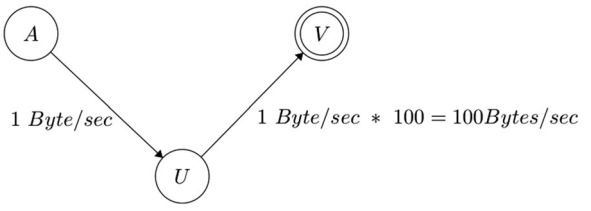

Figure 1. Illustration of a reflective DDoS attack where the attacker A uses a UDP amplifier with an amplification factor of 100 as a reflector. The attacker sends a message to U, with a spoofed IP address that makes it look like it was the victim V that sent the message, resulting in the victim V getting the amplified response from U.

Volumetric attacks can be achieved by multiple means. In this paper though, we will further narrow down our perspective onto a subset of volumetric attacks: UDP amplification attacks,[7] also often called 'reflected DDoS attacks' (rDDoS).

This type of attack is performed by spoofing the *From* address of a packet using a victim's IP address, and sending large numbers of messages to devices that will reply (hereafter called reflectors). This is roughly similar to signing your enemies up for tons of mail-order catalogues. If enough of these reflectors reply, then the victim is overwhelmed.

One important aspect of these reflection DDoS is the 'amplification factor (AF)' of a specific protocol: essentially a simple message to a reflector will result in a large answer from that reflector to the (faked) *From* internet protocol (IP) address. If 1 byte of data arriving at a reflector results in a reflector sending back 100 bytes to the (faked) *From* address, then we would say the AF is 100, as illustrated in Figure 1.

From now on, we will *only focus on UDP-based rDDoS attacks* and – for the sake of simplicity – just call them 'DDoS attacks' (well knowing that this nomenclature is not correct and that other types of DDoS attacks exist, which from now on will be out of scope for this paper).

On the distribution of bandwidth and policy implications

Policy-makers trying to remediate global DDoS potential should primarily focus on high capacity bandwidth UDP amplifiers to reduce destructive DDoS capability. High bandwidth countries or ASNs are the larger contributors. This goes against the naive assumption that we should focus on countries or organisations with the highest counts of reflectors.[8]

The surprising insight we explore below is that increasing the throughput capacity[9] is the primary contributing factor, not, as was previously assumed, the number of reflectors themselves.

The reason for this is simple: the *impact an amplifier can have in a DDoS attack is constrained by its upstream capacity*. For details, see the mathematical description later on in this paper. This fact, combined with the fact that bandwidth capacity on the internet is very unequally distributed, results in the clear policy implication that data centres or other high capacity networks need to be addressed first when cleaning up the DDoS potential (under the very reasonable assumption that we only have certain limited

resources to clean up DDoS amplifiers). *Data centres will be the most cost-effective places for remediation and clean-up efforts*.

How is bandwidth distributed across the internet?

In fact, we can observe the regular scale-free (CAIDA), long-tail distribution of bandwidth capacity. An example of this can be seen in the study by RIPE NCC and M-Lab[10] on bandwidth capacity which clearly shows the long-tail distribution of internet bandwidth. While data centres have gigabit-per-second (Gb/s) connectivity, usual digital subscriber line (DSL) households are in the single-digit megabit-per-second (Mb/s) upstream capacity (or even less). Even worse, mobile phones are further below that rate. The limiting factor in DDoS amplification attacks is the *upstream* bandwidth of the UDP amplifier.

See Figure 2 for an example of the M-Lab data for Russia in the time window 2016-05-10 21:59:00 UTC to 2017-05-10 21:59:00 UTC.

Distribution of bandwidth in Russia between the 10th of May 2016 and the 10th of May 2017 (source: M-Lab, RIPE NCC).

Therefore as we can show with a simple example below (and expand more mathematically in later sections), data centres are a far bigger contributor to the DDoS potential than a few low bandwidth household IoT devices. In our example, the data centre contributes 100 times more to the Global DDoS potential than the DSL devices, despite the fact it only has a tenth of the reflectors (Table 1)!

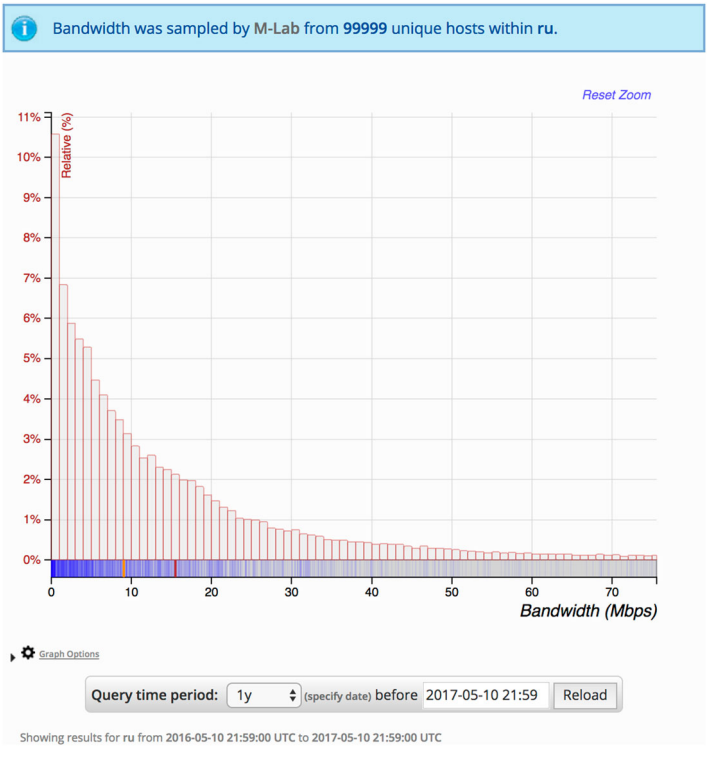

Figure 2. Distribution of bandwidth in Russia between the 10th of May 2016 and the 10th of May 2017. Source: M-Lab, RIPE NCC.

Table 1. A hypothetical example of bandwidth normalised potential.

Location	Number of reflectors	Download	Upload	Potential
Data centre	10	1 Gb/s	1 Gb/s	10
Households	100	10 Mb/s	1 Mb/s	0.1

Given the example above, it should be easy to see that if a country has a low upstream bandwidth, then their DDoS risk to others is limited primarily by that factor, not the count or AF. However, if they are planning on improving bandwidth in the near future it would be wise to also do clean-ups of reflectors at roughly the same time, before they become a problem for others on the internet.

To further progress this metric and approach (beyond simply sourcing more accurate data or extending the methodology to IP version 6 (IPv6)), it would be wise for policy-makers to examine the economics and DDoS.

The first paragraph of 'policy implications' should have a high-level, non-technical conclusion for policy-makers, such as upstream capacity is the key enabler of reflected DDoS attacks. Large data centres have much bigger potential of DDoS contribution than households. The relatively lower number of large data centres, their relatively high resources and technical ability should enable real progress to be made. Great advice exists on how to reconfigure key protocols to make them less capable of reflection and amplification, yet is rarely implemented. Our research findings suggest that the organisations with high capacity upstream could play an important role in reducing the capacity for destructive DDoS attacks affecting the entire internet.

Note too, how heavily the developed world has a higher potential impact on others when we normalise by bandwidth, than when we look at raw reflector counts multiplied by AF. It is primarily bandwidth that constrains the reflective power of servers, and when we add bandwidth we add harmful potential in proportion to the reflector counts. We may think of the new folks of the internet as more harmful in their misconfiguration of devices, but it is those of us with high upload speeds that should be most vigilant in reducing our pollution of the internet via reflectors.

When we first count devices, it is simply a census. This may be how we have done things previously, but this does not give us a true severity or total potential DDoS capacity. It ignores both the amplification and the constraint on bandwidth.

Interestingly though, if high bandwidth and reflectors are present, attackers will benefit. Thus, severity begins to correlate with likelihood of traffic emission over time, as attackers discover 'where the amplification lives'. Over time, DDoS events will show the distribution that attackers uncover, and we should not be surprised if what they uncover correlates with the reflectors that have highest upstream capacities. In other words, expect DDoS reflectors to be most abused where the upload bandwidth is high, not where the largest counts of reflectors are.

Now, let us discuss a more subtle point of policy even than the constrained capacity. DDoS events are multi-causal, with vendor, deployer and attackers all partially to blame for the harms. The vendor ships a device that is easily or by default misconfigured. The deployer does not employ *caveat emptor* or clean up the mess, and the attacker triggers the harm through actively generating malicious traffic.

Thus, DDoS is a collective action problem, and these refinements to our metrics are very much needed to tackle the measurement aspects of that problem and coordinate our

collective action. Notably, policy-makers should also take a look at CAIDA's spoofer project.[11] As measurement improves, so too should policy-makers seek to learn their lessons from the data. Policy needs to shift in the face of new evidence and adapt to the most effective (and cost-effective) mitigation strategies.

Finally, one side note, that the economics of DDoS should be disambiguated fivefold:

1. There are costs to the attackers to scan and send malicious traffic
   a. machine cost
   b. bandwidth cost
   c. development cost
2. There are costs to the reflector hosts in degraded bandwidth and service
   a. bandwidth
   b. service degradation
3. There are costs to the victim in service degradation and loss
   a. bandwidth
   b. service degradation
   c. mitigation
   d. business interruption
4. There are costs to the transporting networks who have to carry traffic ('carriers')
   a. mitigation costs
   b. capacity loss
5. There is revenue for the (IoT) vendor with occasional reputation loss
   a. The vendor of a vulnerable device benefits from device profits, but suffers few of the costs (except reputation in rare cases)

Four types of protocols

In reflected volumetric DDoS, there are a couple of protocols known to produce as many problems as UDP reflectors.[12] Out of these the OpenNTP project (the data source for the authors) chose to measure four particularly well-known ones (see below). We have a reasonable sense of their respective AFs, and a strong history of their impact on DDoS events. We focus primarily on these services, but the technique for estimating DDoS potential is described later in the paper.

1. Network time protocol (NTP) – used for time servers
2. Open recursive domain name system (DNS) – used for the name-to-IP-address mapping service of the internet
3. Simple service discovery protocol (SSDP) – used for 'universal plug and play' – mostly in consumer internet setups
4. Simple network management protocol (SNMP) – usually used for monitoring services and devices on the internet.

Reflectors

These four protocols have services that can be configured poorly and contribute heavily as reflectors and amplifiers online. Great advice[13] already exists on how to reconfigure them

to be less capable of reflection and amplification, yet frustratingly, this advice is rarely followed.

What are we measuring?

We are interested in measuring the severity of harm, if all reflectors in IP version 4 IPv4 were harnessed. This is a unique contribution, as most papers empirically measure DDoS event sizes, but ignore how big it *could* be. Interestingly, we discovered there are sources of empirical data that could help us predict how big a perfect IPv4 reflected DDoS could be.

By combining the number of misconfigured servers scaling it by the AF, and constraining it by the available bandwidth of the reflectors, we get a DDoS potential for each service. This illustrates the malicious potential that a perfect adversary could harness. Finally, we produce a stacked graph showing these risks combined and tracking them to the best of our ability over time.

How did we measure it?

The OpenNTP project developed a scanning technique similar to *zmap* (Durumeric, Wustrow, and Halderman 2013) which periodically measures the whole internet on a weekly basis and produces a data set of IP addresses which are vulnerable to be misused as amplifiers. The authors worked with the CyberGreen project[14] to take this data set and aggregate it per country and network (ASN). This results in the count of amplifiers per country or ASN. The CyberGreen project makes these aggregated data sets freely available on their web page.

Now we can further calculate the impact a country or ASN has (Figure 3):

In words: the DDoS impact of a country c_i is the sum over all devices in country i which are vulnerable to the protocol-amplification j. (Note: $j = 1 \ldots 4$ in our case: DNS, NTP, SSDP, SNMP – four protocols). $AF(j)$ is the AF of protocol j.

Example: The risk 'Open NTP' has an average AF of 557. Source: https://www.us-cert.gov/ncas/alerts/TA14-017A (Table 2).

The geolocation of IP addresses to countries needs to happen in a consistent way which also respects historic IP-to-country assignments (for example, via MaxMind geoip[15]). Specifically, it is important to get the geolocation information at the same time as the

$$impact_{c_i} = \sum_{j=1}^{n} devices_j(c_i) * AF(j)$$

Figure 3. Raw amplified DDoS potential.

Table 2. Protocol specific amplification factors.

Protocol	AF	Comment
NTP	556.9	
DNS	~41	Can vary 28–54
SSDP	30.8	
SNMP	6.3	

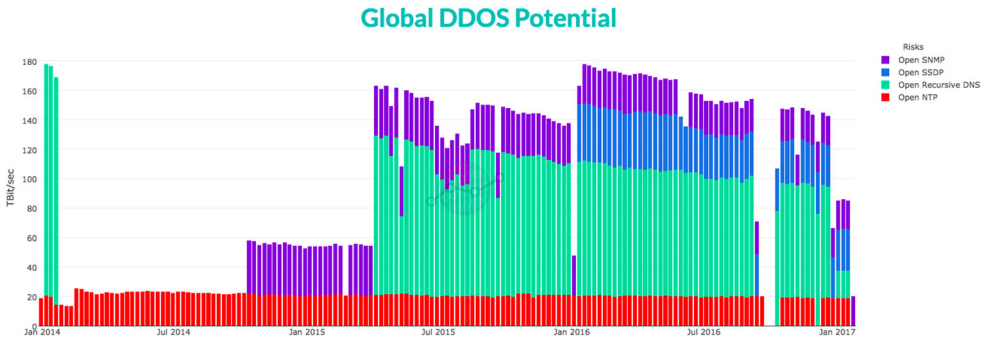

Figure 4. Graph of raw potential over time.

scans for reflectors, or the geolocation is out of date, with respect to the scans. This is beyond the scope of the paper, but is one of the benefits of using CyberGreen data.

Once this is finished you can visualise the sum impacts of all countries and arrive at a global potential of our four protocols across the IPv4 space. On the CyberGreen website they do this with a stacked graph, so that they may understand the proportionality of the contribution of each protocol. However, this is not currently constrained by bandwidth which is why we have written this paper. They may choose to add this normalisation by bandwidth in a future version of the metric (Figure 4).

Global DDoS potential visualisation: This visualisation is based on metrics in CyberGreen 2.1. Note the missing data for the risk 'Open recursive DNS' and 'SSDP' during certain times. This is why we averaged the counts per country across the year 2016, the year with the most robust data (Figure 5).

The truly interesting thing is that bandwidth increases are slowly revealing a lurking titan of rDDoS potential, which is currently only 0.0038% of the summed and amplified reflectors (Table 3)!

In other words, the more bandwidth we add upstream of the reflectors, the more their AFs will matter, and the greater problems we give ourselves.

A naive assumption

The simple idea of multiplying the risk by some constant AF is tempting but it ignores an important constraint of real-world networks: *the upstream of an amplifier is limited*.

Let us look at a simplistic model of an amplification and reflection attack (Figure 6):

Attacker A sends x Bits/sec via a U (UDP amplifier), which then amplifies the data to the victim V by an *AF*. However, the upstream U has a hard capacity limit. Strictly speaking it has both a hard limit and an empirical median which might vary by usage, or indeed other DDoS attacks.

To correct for this constraint in bandwidth, we propose to fix the capped upstream issue by adding a constraint on the amplified count in the CyberGreen application programming interface (API) to account for this capacity limitation, as demonstrated in the equation below (Figure 7).

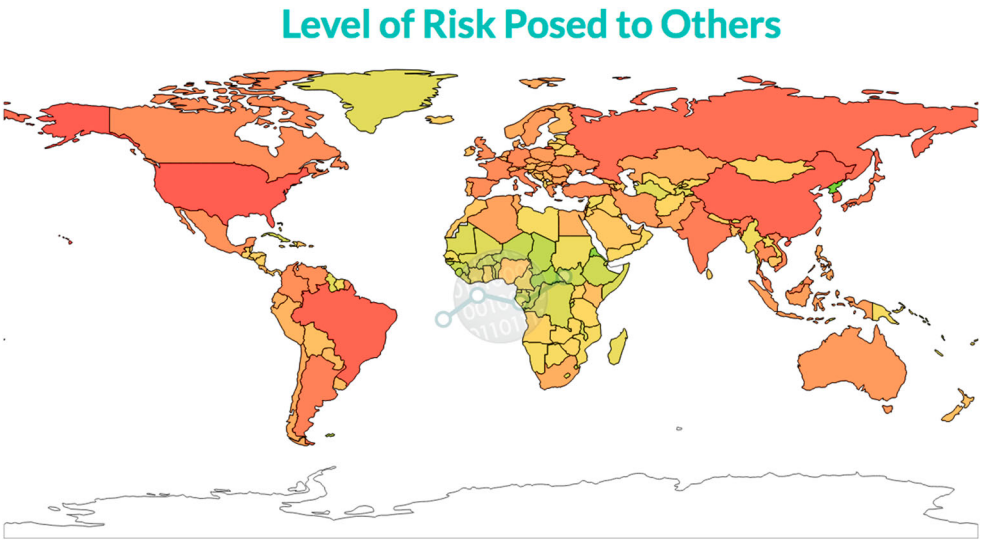

Figure 5. Map of raw potential.

Table 3. Percentage of rDDOS potential revealed by bandwidth.

Raw potential summed	28055.1568345 Tib
Bandwidth constrained potential summed	108.495838045 Tib
Bandwidth revealed percentage	0.0038

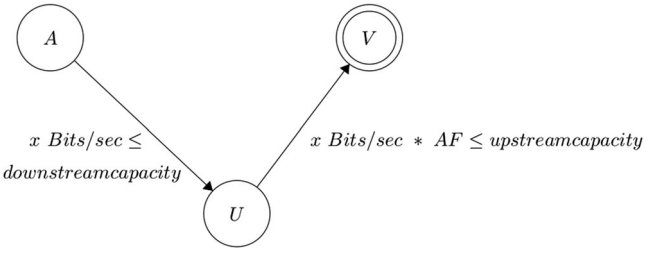

Figure 6. An example of why bandwidth constrains rDDoS potential.

$$impact_{c_i}[\mathbf{MBit/sec}] = \sum_{j=1}^{n} devices_j(c_i) * \min\left(US(c_i); AF(devices_j) * DS(c_i)\right)$$

Figure 7. Equation for bandwidth normalised potential.

Where $US(c_i)$ is the median upstream for a country c_i and $DS(c_i)$ is the median downstream. Also, note that *US* and *DS* are seen as from the perspective of the device, not of the attacker.

We can further refine this metric since there are a couple of data sets which give us an upstream data for individual IP addresses, netblocks, ASNs or countries. M-Lab is one such example data set, which contains empirical data on speed tests conducted worldwide for more than two years.

M-Lab plans to release an API which will make it possible to refine this estimate on a per ASN level, from empirical sources, though for the purposes of this paper, we have used Big Query to produce country-based medians for the month of March (Tariq, Hong, and Khee 2006). We hope other interested scholars will refine this methodology further by computing this on a per ASN median, and progressing the art of rDDoS empirical research. Of course, this methodology is also a victim of the quality of data, and further refinements in data science in internet cartography will produce better and better estimates. Let us not digress from our intriguing results, though.

With the M-Lab data we were able to make a weighted sum over all individual IPs of amplifiers and their respective median upstreams and median downstreams for the year 2016. Note that this then basically becomes the sum of the median upstream speeds of the vulnerable IP addresses (in the majority of cases).

A brief comment on data quality and future research directions

Both CyberGreen and M-Lab conduct internet-wide research, and freely acknowledge the data quality issues that abound in this space. As per these organisations, these issues can exist both because real networks are complex, and scanning and internet cartography are complicated, but also because consistency of collection and temporal data quality are issues.

However, the authors are accustomed to vetting these kinds of data quality issues and chose both CyberGreen and M-Lab for their *transparency* on these matters. These two organisations can and will answer questions about anomalies in their collection of data. Future researchers repeating our methodology may choose to use other data such as that offered by CAIDA (Toponce 2009), and we encourage further debate about such matters.

We believe this is fertile ground for further research, and here offer a few proposals of directions we believe to be promising. We know that this potential is constrained by the edge network, but we have not looked at how it might be constrained at the core network. This could segment the potential and mean that there are islands that can experience heavy DDoS but cannot export it because the max-flow of the core network cannot sustain it.

This or other hidden variables may mean our number is not correct, and we welcome empirical evidence to challenge and improve it. Anecdotally, by collecting information anonymously from 'booter and stresser' operators, we hear that they hit limits of about 15 Tb/s, though these claims are not verifiable.

Some other directions of research would be to answer what proportion of this is IoT devices. It might be reasonable to estimate this as a proportion initially, but in the future it should also be possible to identify devices by various protocols and fingerprinting techniques. The challenges in this direction will be keeping up with the pace of IoT development, and the proportionality will change quickly, especially in IPv6.

It has been posited that IoT is a contributor to the growth of DDoS, but tracking this over time will tell. Of course we would be remiss not to mention that some IPv6 analysis of a similar approach would be a welcome addition, and might begin by estimating proportionality of protocol traffic across utilised IPv6 space. Alternatively, using sinkhole data (routable but unused address space for recording traffic) for IPv6 might give an estimate of the proportions of SMTP/SNMP/DNS/NTP protocols seen there. There are also differences with IPv6 that should be studied and respected too, such as different AFs and wholly new protocols. Finally, we believe that this DDoS potential will show significant temporal variation, and thus tracking it and its changes over time will be fruitful for research purposes and greater understanding of the internet.

Nonetheless this paper represents a first proof of concept that it is entirely possible to calculate a global DDoS potential, which to our knowledge has never been achieved before. The paper also demonstrates that focused intervention on high bandwidth organisations might be very fruitful for reducing this capacity.

Conclusion

Due to the continued growth of the internet, more people are hosting devices that can be used to harm others with DDoS attacks. While some countries are tackling it well, others are getting worse. Counter-intuitively, countries or ASNs who focus on increasing bandwidth increase the severity of their risk to others, which in turn increases their likelihood of being used (since attackers seek strong reflectors). Thus, mitigations are important when increasing network capacity, to keep from rising in the ranks of internet pollution from DDoS.

The primary contribution of this paper is a useful, extensible measurement of DDoS capacity distributed throughout IPv4 space. The method is extensible to IPv6, or to other reflector protocols. It can also be significantly improved with better data. As a consequence of our work, we can see that the potential harm of DDoS already greatly outstrips current defensive capabilities as measured in Tb/s.

In conclusion, we estimate that today's IPv4 internet is capable of at least 108.49 Tb/s in DDoS capacity. However, we also note that this is subject to significant temporal variations. One interesting way to verify our claims would be to ask DDoS mitigation companies to verify if the total simultaneous DDoS attacks sum to more or less than our estimate. In other words, in any given second multiple DDoS events (within 2016) are subsets to the potential we have described here. Do the sum of distinct witnessed DDoS events, approximate the numbers we have here? What proportion of the potential is utilised? Do witnessed distinct events exceed or invalidate our claims in this paper? What is the variance of witnessed events, compared to the variance of max potential? Could a ranking of DDoS events country contributions be compared to our rank list of potential? Would a rank correlation exist between potential and witnessed events? We believe these are exciting research questions to pursue, towards an actuarial comparison to an rDDoS model that could be created with this metric. Putting this in simple terms, we believe DDoS could one day be modelled mathematically, similar to how hurricanes are modelled today.

To be absolutely clear about our methodology (in aid of further research), this result was achieved by using the average reflector counts across all of 2016 for each country, and then constrained by the median downstream to the reflector or the upstream from it,

whichever is smaller (for each country). The countries are then summed to get the final numbers.

This gives us an interesting single result which can be replicated year on year to gain a better understanding of DDoS attacks. Of course we could also calculate quartiles to understand more of the temporal variance (in both throughput and reflector counts). Tracking this over time will lead to many more policy realisations, and ultimately aid management and mitigation of the global DDoS problem.

By contrast with the calculated potential, the largest DDoS event we have seen is 1.2 Tb/s reported by a single target: Dyn (Hilton 2016). In other words, the worst we have seen is an order of magnitude smaller than the untapped malicious potential of the internet. Likewise, our defenses are far outmatched by the potential, something the larger DDoS protection companies might consider for further research.

Notes

1. http://openntpproject.org/.
2. https://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/.
3. https://en.wikipedia.org/wiki/Network_Time_Protocol.
4. https://en.wikipedia.org/wiki/Domain_Name_System.
5. https://en.wikipedia.org/wiki/Simple_Service_Discovery_Protocol.
6. https://en.wikipedia.org/wiki/Simple_Network_Management_Protocol.
7. https://en.wikipedia.org/wiki/Denial-of-service_attack#Reflected_.2F_spoofed_attack.
8. As the CyberGreen project does.
9. Especially the upstream capacity.
10. https://labs.ripe.net/Members/vastur/visualising-bandwidth-capacity-in-ripestat-using-m-lab-data.
11. https://www.caida.org/projects/spoofer/.
12. https://www.us-cert.gov/ncas/alerts/TA14-017A.↩
13. https://www.us-cert.gov/ncas/alerts/TA14-017A.↩
14. https://stats.cybergreen.net.↩
15. https://maxmind.com.↩

Acknowledgements

This paper would not be possible without a lot of knowledgeable people – especially Jared Mauch of the OpenNTP project and in and around the CyberGreen project, in particular Yurie Ito, Joe St. Sauver, Paul Vixie, Moto-san Kawasaki and Dr Dan Geer amongst many others. Without the continuous support and – most importantly – meetings and discussions with RIPE NCC folks, the authors would have been lost. Special thanks to Mirjam Kuehne, Emile Aben, Christian Teuschel and Anand Buddhev. The authors would also like to thank M-Lab support team for rapid assistance in using Big Query. In particular Collin Anderson, Chris Ritzo and Georgia Bullien. Network Nerds of the Finest Kind.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Eireann Leverett is an Open Web Fellow at Privacy International, a Senior Risk Researcher at Cambridge Centre of Risk Studies, and an entrepreneur at Concinnity Risks. He enjoys the intersections of risk, hacking, metrics, economics, and liability.

Aaron Kaplan is a Unix user and programmer since 4.3BSD-Lite / FreeBSD 1.0. He has worked for major telecoms, IBM, ESA, banks and critical infrastructure industries since 1997. He is part of a team responsible for running the national CERT - CERT.at. There, he focuses on incident handling automation on a countrywide scale. He is on the board of directors of FIRST.org since 2014.

ORCID

Eireann Leverett http://orcid.org/0000-0001-6586-7359

References

Asosheh, Abbass, and Naghmeh Ramezani. 2008. "A Comprehensive Taxonomy of DDoS Attacks and Defense Mechanism Applying in a Smart Classification." *WSEAS Transactions on Computers* 7 (4): 281–290.

Barabási, Albert-László. 2009. "Scale-free Networks: A Decade and Beyond." *Science* 325 (5939). http://science.sciencemag.org/content/325/5939/412.

CAIDA. Data sets, 2017-06-16. https://www.caida.org/data/overview/.

Durumeric, Zakir, Eric Wustrow, and J. Alex Halderman. 2013. "ZMap: Fast Internet-wide Scanning and Its Security Applications." *Usenix Security* 8: 47–53.

Hilton, Scott. 2016. "Dyn Analysis Summary of Friday October 21 Attack." https://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/.

M-Lab. NDT data set, 2017-04-12. https://measurementlab.net/tools/ndt.

Mirkovic, Jelena, and Peter Reiher. 2004. "A Taxonomy of DDoS Attack and DDoS Defense Mechanisms." *ACM SIGCOMM Computer Communication Review* 34 (2): 39–53.

Tariq, Usman, M. Hang, and K. Lhee. 2006. "A Comprehensive Categorization of DDoS Attack and DDoS Defense Techniques." *ADMA LNAI* 4093. https://link.springer.com/chapter/10.1007/11811305_112.

Toponce, Aaron. 2009. "The Sheer Size of IPv6." https://pthree.org/2009/03/08/the-sheer-size-of-ipv6/.